



Energy-Efficient Memory Designs based on Partial WCET Analysis and Variable Retention-Time NVMs

Rabab Bouziane, <u>Erven Rohou</u>, Abdoulaye Gamatié RTNS, Chasseneuil du Poitou, Oct 2018

Acknowledgments

- Part of Rabab Bouziane's PhD
 - co-advised by Abdoulaye Gamatié and myself

- Partially founded par ANR
 - Continuum
 - ANR-15-CE25-0007-01
 - coordinated by Abdoulaye Gamatié, LIRMM







Power Consumption and Non Volatile Memories

- Static power is becoming dominant
- Non Volatile Memories provide a solution
 - "infinite" retention
 - no leakage, quasi-zero static power

- Very active research domain
- Various technologies
 - PCRAM, STT-RAM, RRAM, FRAM...

Salient Characteristics

Feature	SRAM	DRAM	Disk	Flash	STT-RAM	PCRAM	RRAM
Cell size	>100F ²	6-8F ²	2/3F ²	4-5F ²	37F ²	8-16F ²	>5F ²
Read latency	<10 ns	10-50 ns	8.5ms	25 µs	<10 ns	48 ns	<10 ns
Write latency	<10 ns	10-60 ns	9.5 ms	200 µs	12.5 ns	40-150 ns	10 ns
Energy per bit access	>1 pJ	2 pJ	100-1000 mJ	10 nJ	2 pJ	100 pJ	0.02 pJ
Leakage power	High	Medium	High	Low	Low	Low	Low
Endurance	> 1 0 ¹⁵	>10 ¹⁵	>10 ¹⁵	104	> 10 ¹⁵	10 ⁵ -10 ⁹	10 ⁵ -10 ¹¹
Non volatlitly	No	No	Yes	Yes	Yes	Yes	Yes
Scalability	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Jishen Zhao, Cong Xu, Ping Chi, and Yuan Xie. "Memory and Storage System Design with Nonvolatile Memory Technologies". In: IPSJ Transactions on System LSI Design Methodology (2015).



Should the compiler care?

Compilers eliminate memory accesses, anyhow!

- Still:
 - R/W asymmetry
 - silent stores elimination
 - Variable retention time
 - this work

Bouziane, R., Rohou, E. and Gamatié, A., 2018. Compile-Time Silent-Store Elimination for Energy Efficiency: an Analytic Evaluation for Non-Volatile Cache Memory. In RAPIDO'18.



Non volatile?

- Meaning:
 - volatile with long retention time
 - typically in years
- But
 - just a parameter
 - can be adjusted

Retention time	Read energy (nJ)	Write energy (nJ)
4.27 yr	0.085	1.916
3.24 s	0.083	0.932
26.5 µs	0.081	0.347

Sun, Z., Bi, X., Li, H. H., Wong, W. F., Ong, Z. L., Zhu, X., & Wu, W. (2011). Multi retention level STT-RAM cache designs with a dynamic refresh scheme. In *MICRO*.

Retention time	Read energy (nJ)	Write energy (nJ)
10 yr	0.233	0.601
10 ms	0.233	0.269

Khoshavi, N., Chen, X., Wang, J. and DeMara, R.F., 2016. Read-tuned STT-RAM and eDRAM cache hierarchies for throughput and energy enhancement. arXiv preprint arXiv:1607.08086.

Opportunity

Lifetimes are not equal



#define N 10 int main(void)

int a, b, i; a = N;

Our Approach

- Identify lifetimes
- Compute worst-case duration of each lifetime
- Design multi-bank NVM
- Assign each lifetime to most appropriate bank

Identification of lifetimes

- Definition: a value is alive from its definition to the last use (before it gets rewritten)
 - Note: one variable may have multiple lifes

- Standard compiler analysis
 - dataflow analysis
 - def-use chains



Heptane: Static Tool for Computing WCET Estimates

- High-level analysis
 - CFG extraction
 - context-sensitive
- Low-level analysis
 - machine model
- Build ILP problem
 - CPLEX, lp_solve, Gurobi...

D. Hardy, B. Rouxel, I. Puaut. The Heptane Static Worst-Case Execution Time Estimation Tool. 17th International Workshop on Worst-Case Execution Time Analysis (WCET 2017).



Worst-case duration of lifetimes

- Introduce δ-WCET
 - WCET between two given basic blocks
 - much tighter for subgraphs
 - context-sensitive
- Based on DFS walk
 - special handling of *maxiter* annotation
- Extension of Heptane



Multi-bank NVM

- Possible exploitation: ISA extension
- Code rewrite



Experimental setup

Hardware

- 40 MHz
 - 1-cycle memory access
- no cache
 - embedded/IoT devices
 - simpler analysis

Software

- extended Heptane
- Mälardalen benchmarks
 - increased execution time
 - gcc -00



Worst-Case Lifetimes

- Distribution confirms optimization potential
 - more data in paper



Memory Subsystem Energy Savings



Conclusion and Perspectives

- Opportunity to reduce energy consumption due to:
 - Variable retention time in memory
 - WCET estimates

- Next
 - compiler-level assignment
 - support caches
 - data layout

E. Rohou - RTNS 2018



Thank you!

E. Rohou - RTNS 2018



Ínría



